Machine Learning MCQ Questions and Answer PDF

1. Type of matrix decomposition model is_____

1. predictive model
2. descriptive model
3. logical model
4. None

Answer: descriptive model

2. PCA is_____

1. backward feature selection
2. forward feature selection
3. feature extraction
4. None of these

Answer: feature extraction

3. Supervised learning and unsupervised clustering both require which is correct according to the statement.

1. input attribute
2. hidden attribute
3. output attribute
4. categorical attribute

Answer: input attribute.

4. Following are the types of supervised learning_____

1. regression
2. classification
3. subgroup discovery
4. All of above

Answer: All of above

5. A feature F1 can take certain value: A, B, C, D, E, & F and represents grade of students from a college. Here feature type is_____

1. ordinal
2. nominal
3. categorical
4. Boolean

Answer: ordinal

6. Following is powerful distance metrics used by Geometric model_____

1. Manhattan distance
2. Euclidean distance
3. All of above
4. None of above

Answer: All of above

7. The output of training process in machine learning is_____

1. machine learning algorithm
2. machine learning model
3. null
4. accuracy

Answer: machine learning model

8. Which of the following is a good test dataset characteristic?

1. is representative of the dataset as a whole
2. large enough to yield meaningful results
3. All of above
4. None of above

Answer: All of above

9. Which of the following techniques would perform better for reducing dimensions of a data set?

1. removing columns which have high variance in data
2. removing columns which have too many missing value
3. removing columns with dissimilar data trends
4. None of the above

Answer: removing columns which have too many missing values

10. What characterize is hyper plane in geometrical model of machine learning?

1. a plane with 1 dimensional fewer than number of input attributes
2. a plane with 1 dimensional more than number of input attributes
3. a plane with 2 dimensional more than number of input attributes
4. a plane with 2 dimensional fewer than number of input attributes

Answer: a plane with 2 dimensional fewer than number of input attributes

11. You are given reviews of few Netflix series marked as positive, negative and neutral. Classifying reviews of a new Netflix series is an example of_____

1. unsupervised learning
2. semi supervised learning
3. supervised learning

4. reinforcement learning

Answer: supervised learning

12. Like the probabilistic view, the _____ view allows us to associate a probability of membership with each classification

1. deductive
2. exampler
3. classical
4. inductive

Answer: inductive

13. The problem of finding hidden structure in unlabeled data is called_____

1. unsupervised learning
2. reinforcement learning
3. supervised learning
4. None

Answer: unsupervised learning

14. If machine learning model output involves target variable then that model is called as_____

1. predictive model
2. descriptive model
3. reinforcement learning
4. All of above

Answer: predictive model

15. Database query is used to uncover this type of knowledge.

1. hidden
2. shallow
3. deep
4. multidimensional

Answer: multidimensional

16. Data used to build a data mining model.

1. training data
2. hidden data
3. test data
4. validation data

Answer: training data

17. Application of machine learning methods to large databases is called_____

1. big data computing
2. artificial intelligence
3. data mining
4. internet of things

Answer: data mining

18. Which learning Requires Self-Assessment to identify patterns within data?

1. supervised learning
2. unsupervised learning
3. semi supervised learning
4. reinforced learning

Answer: unsupervised learning

19. In simple term, machine learning is_____

1. prediction to answer a query
2. training based on historical data
3. All of above
4. None of above

Answer: All of above

20. Of the Following Examples, Which would you address using an supervised learning Algorithm?

1. given a set of news articles found on the web, group them into set of articles about the same story
2. given email labeled as spam or not spam, learn a spam filter
3. given a database of customer data, automatically discover market segments and group customers into different market segments
4. find the patterns in market basket analysis

Answer: given email labeled as spam or not spam, learn a spam filter

21. If machine learning model output doesn't involves target variable then that model is called as_____

1. predictive model
2. descriptive model
3. reinforcement learning
4. all of the above

Answer: descriptive model

22. In what type of learning labelled training data is used_____

1. supervised learning
2. unsupervised learning
3. reinforcement learning
4. active learning

Answer: supervised learning

23. In the example of predicting number of babies based on stork's population ,Number of babies is_____

1. feature
2. observation
3. outcome
4. attribute

Answer: outcome

24. Following are the descriptive models_____

1. classification
2. clustering
3. association rule
4. Both 1 and 2

Answer: Both 1 and 2

25. In following type of feature selection method we start with empty feature set_____

1. backward feature selection
2. forward feature selection
3. All of above
4. None of above

Answer: forward feature selection

26. A person trained to interact with a human expert in order to capture their knowledge.

1. knowledge developer
2. knowledge programmer
3. knowledge engineer
4. knowledge extractor

Answer: knowledge extractor

27. What characterize unlabeled examples in machine learning_____

1. there is plenty of confusing knowledge
2. there is prior knowledge
3. there is no confusing knowledge
4. there is no prior knowledge

Answer: there is plenty of confusing knowledge

28. What does dimensionality reduction reduce?

1. collinearity
2. stochastic
3. entropy
4. performance

Answer: collinearity

29. Some telecommunication company wants to segment their customers into distinct groups ,this is an example of_____

1. supervised learning
2. unsupervised learning
3. data extraction
4. reinforcement learning

Answer: unsupervised learning

30. Which of the following is the best machine learning method?

1. accuracy
2. scalable
3. fast
4. All of above

Answer: All of above

31. In multiclass classification number of classes must be_____

1. equals to two
2. less than two
3. greater than two
4. None

Answer: greater than two

32. Which of the following can only be used when training data are linearly separable?

1. linear logistic regression
2. linear hard-margin svm
3. linear soft margin svm
4. parzen windows

Answer: linear hard-margin svm

33. Which of the following can only be used when training data are linearly separable?

1. linear logistic regression
2. linear soft margin svm
3. linear hard-margin svm
4. the centroid method

Answer: linear hard-margin svm

34. You are given seismic data and you want to predict next earthquake , this is an example of_____

1. supervised learning
2. unsupervised learning
3. reinforcement learning
4. dimensionality reduction

Answer: supervised learning

35. Prediction is_____

1. discipline in statistics used to find projections in multidimensional data
2. value entered in database by expert
3. the result of application of specific theory or rule in a specific case
4. independent of data

Answer: the result of application of specific theory or rule in a specific case

36. Impact of high variance on the training set ?

1. under fitting
2. over fitting
3. both under fitting & over fitting
4. depends upon the dataset

Answer: over fitting

37. Which of the following is an example of feature extraction?

1. applying pca to project high dimensional data
2. construction bag of words from an email
3. removing stop words
4. forward selection

Answer: applying pca to project high dimensional data

38. The effectiveness of an SVM depends upon_____

1. kernel parameters
2. selection of kernel
3. soft margin parameter
4. All of the above

Answer: selection of kernel

39. What do you mean by a hard margin?

1. the svm allows very low error in classification
2. the svm allows high amount of error in classification
3. All of above
4. None of above

Answer: the svm allows very low error in classification

40. Which of the following is a reasonable way to select the number of principal components "k"?

1. choose k to be 99% of m (k = 0.99*m, rounded to the nearest integer)
2. choose k to be the smallest value so that at least 99% of the variance is retained
3. choose k to be the largest value so that 99% of the variance is retained
4. use the elbow method

Answer: choose k to be the smallest value so that at least 99% of the variance is retained

41.A student Grade is a variable F1 which takes a value from A,B,C and D. Which of the following is True in the following case?

1. variable f1 is an example of ordinal variable
2. it doesn\t belong to any of the mentioned categories
3. variable f1 is an example of nominal variable
4. it belongs to both ordinal and nominal category

Answer: variable f1 is an example of ordinal variable

42. What is the purpose of the Kernel Trick?

1. To transform the problem from regression to classification
2. To transform the problem from supervised to unsupervised learning.
3. To transform the data from nonlinearly separable to linearly separable
4. All of above

Answer: to transform the data from nonlinearly separable to linearly separable

43. Feature can be used as a_____

1. predictor
2. binary split
3. All of above

4. None of above

Answer: All of above

44. What can be major issue in Leave-One-Out-Cross-Validation (LOOCV)?

1. high variance
2. low variance
3. faster runtime compared to k-fold cross validation
4. slower runtime compared to normal validation

Answer: high variance

45. The cost parameter in the SVM means_____

1. the kernel to be used
2. the trade-off between misclassification and simplicity of the model
3. the number of cross-validations to be made
4. None

Answer: the trade-off between misclassification and simplicity of the model

46. Which of the following evaluation metrics cannot be applied in case of logistic regression output to compare with target?

1. accuracy
2. auc-roc
3. logloss
4. mean-squared-error

Answer: mean-squared-error

47. A measurable property or parameter of the data-set is_____

1. training data
2. test data
3. feature
4. validation data

Answer: feature

48. Support Vector Machine is_____

1. geometric model
2. probabilistic model
3. logical model
4. none

Answer: geometric model

49. Imagine a Newly-Born starts to learn walking. It will try to find a suitable policy to learn walking after repeated falling and getting up. Specify what type of machine learning is best suited?

1. regression
2. means algorithm
3. reinforcement learning
4. None

Answer: reinforcement learning

50. Different learning methods does not include?

1. deduction
2. memorization
3. analogy
4. introduction

Answer: introduction